

# Hybrid Gain Data Assimilation Using Variational Corrections in the Subspace Orthogonal to the Ensemble

CHIH-CHIEN CHANG

*Department of Atmospheric Sciences, National Central University, Jhongli, Taiwan*

STEPHEN G. PENNY

*Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, and  
Physical Sciences Division, NOAA/Earth System Research Laboratory, Boulder, Colorado, and  
RIKEN Center for Computational Science, Kobe, Japan*

SHU-CHIH YANG

*Department of Atmospheric Sciences, National Central University, Jhongli, Taiwan, and  
RIKEN Center for Computational Science, Kobe, Japan*

(Manuscript received 1 May 2019, in final form 16 March 2020)

## ABSTRACT

The viability of a parameterless hybrid data assimilation algorithm is investigated. As an alternative to the traditional hybrid covariance scheme, hybrid gain data assimilation (HGDA) was proposed to blend the gain matrix derived from the variational method and the ensemble-based Kalman filter (EnKF). A previously proposed HGDA algorithm uses a two-step process applying the EnKF with a variational update. The algorithm is modified here to limit the variational correction to the subspace orthogonal to the ensemble perturbation subspace without the use of a hybrid weighting parameter, as the optimization of such a parameter is nontrivial. The modified HGDA algorithm is investigated with a quasigeostrophic (QG) model. Results indicate that when the climatological background error covariance matrix  $\mathbf{B}$  and the observation error covariance  $\mathbf{R}$  are well estimated, state estimates from the parameterless HGDA are more accurate than the parameter-dependent HGDA. The parameterless HGDA not only has potential advantages over the standard HGDA as an online data assimilation algorithm but can also serve as a valuable diagnostic tool for tuning the  $\mathbf{B}$  and  $\mathbf{R}$  matrices. It is also found that in this QG model, the empirically best static  $\mathbf{B}$  matrix for the stand-alone 3DVAR has high variance at larger spatial scales, which degrades the accuracy of the HGDA systems and may not be the best choice for hybrid methods in general. A comparison of defining the orthogonal subspace globally or locally demonstrates that global orthogonality is more advantageous for stabilizing the hybrid system and maintains large-scale balances.

## 1. Introduction

The most effective data assimilation (DA) algorithms used today in operational numerical weather prediction (NWP) form a hybrid combination of Variational (VAR) and ensemble-based methods. Following the pioneering work of Barker (1998), different algorithms have been proposed (Hamill and Snyder 2000; Lorenc 2003; Buehner 2005; Buehner et al. 2010) to realize the

concept of hybrids via blending the climatological background error covariance used for the VAR with a dynamic background error covariance derived from flow-dependent information provided by an ensemble forecast or ensemble Kalman filter (EnKF; Evensen 1994). Hybrid DA has been implemented at operational centers such as the European Centre for Medium-Range Weather Forecasts (ECMWF) (Isaksen et al. 2010) and the National Centers for Environmental Prediction (NCEP) (Kleist and Ide 2015).

Hybrid covariance data assimilation (HCDA) combines the climatological and ensemble-based background error covariance matrices. The climatological background error covariance is typically static and full

---

 Denotes content that is immediately available upon publication as open access.

---

*Corresponding author:* Chih-Chien Chang, changchichien@gmail.com

DOI: 10.1175/MWR-D-19-0128.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) ([www.ametsoc.org/PUBSReuseLicenses](http://www.ametsoc.org/PUBSReuseLicenses)).

rank, while the rank of the ensemble-estimated background error covariance cannot exceed  $k - 1$ , where  $k$  is the ensemble size. From a variational perspective, the flow-dependent information from the ensemble can be introduced to the variational system to improve upon the static background error covariance matrix. The hybrid covariance matrix is typically determined as a weighted sum of the two covariance terms (Hamill and Snyder 2000; Houtekamer and Zhang 2016), and has been extended using alternative algorithms such as the control variable approach (Lorenc 2003).

Forming a hybrid of two error covariance matrices can also be carried out under the EnKF framework. For example, perturbations randomly drawn from the climatological background error covariance can be added to the ensemble perturbations in order to augment the flow-dependent background error covariance with climatological error information. This approach is commonly referred to as additive inflation (Mitchell and Houtekamer 2000) because the sampled information is directly added to the flow-dependent error covariance matrix, though the term “additive inflation” does not necessarily imply a hybrid.

Penny (2014) introduced a hybrid gain data assimilation (HGDA) approach that forms a hybrid combination of the gain matrices that are determined by the EnKF and variational algorithms. Further, Penny (2014) showed that an efficient variant of the HGDA can be implemented without explicitly forming the hybrid gain matrix. This algorithm recenters the ensemble perturbations at a new hybrid mean state and relies on the model dynamics to update the ensemble perturbations during the forecast. Bishop et al. (2017) also proposed a gain form of the ensemble transform Kalman filter (GETKF) using a hybrid gain matrix, which updates both the ensemble mean and the ensemble perturbations during the hybridization analysis step.

The HGDA algorithm has been tested with numerical models with different complexity, including a simple 40-variable Lorenz model (Penny 2014), the NCEP global ocean data assimilation system (GODAS) (Penny et al. 2015), the ECMWF semioperational configuration (Bonavita et al. 2015), and more recently the Canadian Meteorological Centre (CMC) (Houtekamer et al. 2018). In all cases, it was shown that the HGDA improves accuracy compared to the individual EnKF and variational (e.g., 4DVAR or 4DEnVAR) components. These studies have shown that HGDA can 1) stabilize the EnKF system, 2) reduce model bias occurring in the EnKF system, 3) mitigate the misrepresentation of long-range correlation in the covariance matrix caused by localization, and 4) reduce the sensitivity to tuning parameters such as localization and inflation.

Investigations with the Lorenz model focused on the stability of the cycled DA process. The stability of the EnKF relies on the subspace spanned by ensemble members being representative of the unstable-neutral subspace (Bocquet and Carrassi 2017). Results with the Lorenz model show that when the ensemble size is insufficient, using climatological information from the VAR can stabilize the EnKF component in both the HGDA and HCDA approaches by increasing the dimensionality of the solution space beyond that spanned by the ensemble members.

In real-world applications, all DA systems are affected by systematic model errors that lead to model biases (Dee 2005). The presence of such systematic errors can alter the optimal combination weight in a hybrid system (Etherton and Bishop 2004). The EnKF alone is particularly susceptible to systematic model errors because the background error covariance is estimated solely by using an evolved ensemble of biased forecast models (Penny 2017). Because the climatological background error covariance matrix is often constructed using information about forecast errors using many cases spread out over time (Derber and Bouttier 1999; Bannister 2008), it partially considers systematic model errors. The forecast errors are represented by differences between forecasts of different lengths and valid at the same time and this allows the climatological background error partially considering systematic model errors. As demonstrated in the Hybrid-GODAS system (Penny et al. 2015; Penny 2017), the climatological information provided in the global variational correction can remove large-scale temperature and salinity biases arising from the use of localization and an undersampled ensemble.

Although hybrid methods have proven to be effective, the need to optimize the hybrid coefficient remains an unresolved issue (Wang et al. 2013). Both the HCDA and HGDA methods employ a tunable parameter that weights the dynamic and static information supplied by the EnKF and VAR. An empirically best weight can be found based on sensitivity tests for particular cases or long-term investigation of the hybrid system (Storto et al. 2018). For instance, the NCEP Global Data Assimilation System (GDAS) uses 12.5% and 87.5% static and ensemble background error covariance, respectively (Huang and Wang 2018). Under the HCDA framework, Ménétrier and Auligné (2015) demonstrate a variational method to optimize the hybrid covariance weight and the localization parameter simultaneously. In practice, the weight can be a single value or dependent on the wavenumber of the model state (Kleist 2012) or vary vertically (D. Kleist 2015, personal communication). The optimal weight can also

be affected by the observation density (Satterfield et al. 2018). However, tuning an empirical best weight spatially introduces additional degrees of freedom into the hybrid system. Given the difficulties of estimating a proper weight in the hybrids, we were motivated to develop a parameterless hybrid algorithm.

The present study proposes a new algorithm to avoid the use of a combination weight by leveraging the variational correction used in the HGDA algorithm. The HGDA uses the EnKF to reduce the dynamical error structure via a contraction within the ensemble perturbation subspace, while the unrepresented unstable and neutral modes are further constrained using a full-rank static background error covariance matrix using a variational minimization. Focusing on refining the confining the space for correction, Carrassi et al. (2008) used a similar approach with a 3D variational assimilation in the unstable subspace (3DVAR-AUS) to assimilate each observation based on the spatially unstable structure defined by bred vectors (Toth and Kalnay 1993). Both HGDA and 3DVAR-AUS suggest the importance of obtaining the corrections associated with the growing dynamical error modes. During the efficient two-step HGDA algorithm, the solution from the EnKF is determined by the linear basis formed by the analysis ensemble perturbations, while the variational optimization applies an additional correction within the linear space representative of the full system dimension. Assuming the EnKF solution is accurate within the subspace defined by the analysis ensemble perturbations, it is preferred that any additional correction from the VAR should be restricted to the subspace orthogonal to the linear basis defined by the analysis ensemble perturbations.

The concept of applying the orthogonality to form a hybrid DA system has been applied in the reduced rank Kalman filter (Heemink et al. 2001; Petrie and Bannister 2011). The algorithm proposed by Heemink et al. (2001), named as Partially orthogonal ensemble Kalman filter (POEnKF), applies the orthogonality to account for the information lost by the ensemble. The POEnKF algorithm aims to blend the reduced-rank square root filter (RRSQRT; Verlaan and Heemink 1997), which uses the  $p$  leading eigenvectors to form a covariance matrix, and an independent EnKF system with  $k$  randomly generated ensemble members. These two DA systems interact with each other only at the analysis step to form a hybrid covariance. In the POEnKF, the RRSQRT system provides the direction of the leading  $p$  eigenvectors and only the information of the EnKF orthogonal to the subspace spanned by the  $p$  eigenvectors is used as a supplement. In our proposed algorithm, the EnKF first corrects errors in the subspace

spanned by the ensemble members and only the part of the variational update orthogonal to the EnKF analysis perturbations is used as a correction to the EnKF analysis mean.

The variational information can be updated locally or globally to correct the EnKF analysis mean state. It is expected that applying the variational correction globally can constrain the large-scale bias and errors and preserve the dynamical balance of the variational correction. However, updating the variational information at each local region is attractive for computational efficiency. Penny (2014) found that applying the variational information globally leads to better accuracy in the HGDA. In our proposed algorithm, the orthogonal correction from the variational correction can be defined globally and locally. The impact of adopting global orthogonality and local orthogonality is investigated.

In what follows, section 2 provides a detailed methodology about the standard HGDA algorithm and the newly proposed version that limits the VAR correction to subspace orthogonal to the analysis ensemble perturbations. Section 3 introduces the model and the setup of the numerical experiments. Data assimilation results and sensitivity experiments are presented in sections 4 and 5, respectively. Section 6 provides a comparison between global orthogonality and local orthogonality.

## 2. Methodology

### a. Hybrid gain algorithm

The HGDA can be generalized as a hybrid combination of gain matrices derived from arbitrary sources. Penny (2014) proposed a hybrid combination of gain matrices determined via the EnKF and a variational method. The resulting hybrid gain matrix ( $\mathbf{K}$ ) can be expressed as

$$\hat{\mathbf{K}} = c_1 \mathbf{K} + c_2 \mathbf{K}_B + c_3 (\mathbf{K}_B \mathbf{H} \mathbf{K}), \tag{1}$$

where  $\mathbf{K}$  and  $\mathbf{K}_B$  denote the gain matrices from the EnKF and the VAR, respectively;  $\mathbf{K}_B \mathbf{H} \mathbf{K}$  represents potential interactions between two DA systems, where  $\mathbf{H}$  is the linearized observation operator. With an appropriate choice of the constants  $c_1$ ,  $c_2$  and  $c_3$ , we can construct an efficient algorithm for determining  $\hat{\mathbf{K}} \mathbf{d}$ , where  $\mathbf{d}$  is the observation innovation. Penny (2014) proposed two different strategies to conduct the hybrid gain algorithm. Figure 1 shows the flowchart of these two HGDA scenarios: “a” and “b.” Scenario a uses hybrid coefficients  $c_1 = 1$ ,  $c_2 = \alpha$ , and  $c_3 = -\alpha$ , while scenario b uses  $c_1 = (1 - \alpha)$ ,  $c_2 = \alpha$ , and  $c_3 = 0$ .

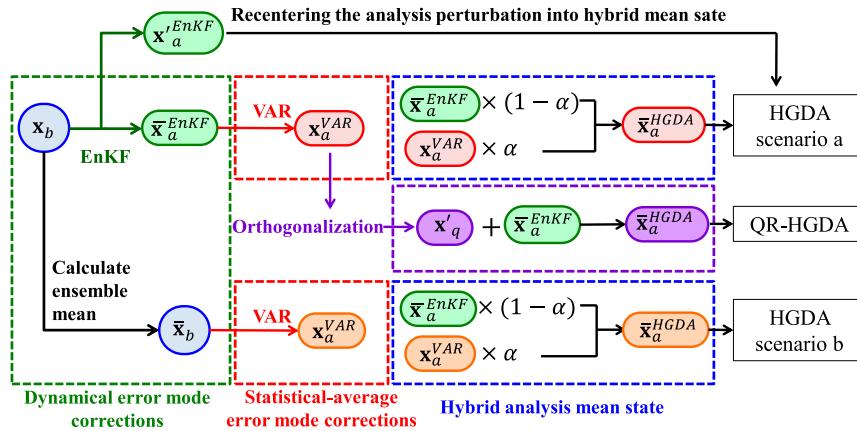


FIG. 1. Flowchart of different versions of HGDA. In the HGDA scenario a, a standard EnKF is executed first and the analysis mean state is then used as the background field of the VAR in the second step update. In the HGDA scenario b, both the EnKF and the VAR use the same background field. After the update process, an empirical optimal combination weight,  $\alpha$ , is given in both scenario a and scenario b to form a hybrid analysis mean state. In the QR-HGDA, the component that is orthogonal to the ensemble subspace is extracted from the VAR’s analysis and used as the second step correction directly. Thus, the QR-HGDA avoids the use of an empirical parameter.

Scenario a requires a two-step sequential update. In the first step, a standard EnKF with  $k$  ensemble members is used to determine the analysis mean state and the analysis ensemble perturbations. The mean vector ( $\bar{\mathbf{x}}_a^{\text{EnKF}}$ ) with a model dimension of  $m$  is given by

$$\bar{\mathbf{x}}_a^{\text{EnKF}} = \bar{\mathbf{x}}_b + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}_b). \tag{2}$$

The perturbation matrix ( $\mathbf{X}_a^{\text{EnKF}}$ ) has a dimension of  $m \times k$ . The background ensemble mean state ( $\bar{\mathbf{x}}_b$ ) has a dimension of  $m$ , and the operator  $\mathbf{H}$ , transforming the model state variables from the model space to the observational space, has a dimension of  $l \times m$ . The second step corrects the analysis mean state of the EnKF by supplying it as the background field of the VAR. The resulting cost functional for the variational minimization is

$$J(\mathbf{x}) = (\mathbf{x}_a^{\text{VAR}} - \bar{\mathbf{x}}_a^{\text{EnKF}})\mathbf{B}^{-1}(\mathbf{x}_a^{\text{VAR}} - \bar{\mathbf{x}}_a^{\text{EnKF}}) + (\mathbf{y}^o - \mathbf{H}\mathbf{x}_a^{\text{VAR}})\mathbf{R}^{-1}(\mathbf{y}^o - \mathbf{H}\mathbf{x}_a^{\text{VAR}}). \tag{3}$$

To minimize this functional, the gradient of  $J$  [Eq. (3)] is set to zero so that the VAR analysis can be derived as

$$\mathbf{x}_a^{\text{VAR}} = \bar{\mathbf{x}}_a^{\text{EnKF}} + \mathbf{K}_B(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}_a^{\text{EnKF}}). \tag{4}$$

The solutions of the VAR analysis and the EnKF analysis mean state are then blended to form the hybrid analysis mean state:

$$\bar{\mathbf{x}}_a^{\text{HGDA}} = (1 - \alpha)\bar{\mathbf{x}}_a^{\text{EnKF}} + \alpha\mathbf{x}_a^{\text{VAR}}. \tag{5}$$

The tunable parameter  $\alpha$  reflects the expected accuracy of component systems. The ensemble analysis perturbations of the EnKF are then recentered at the hybrid analysis mean state:

$$\mathbf{X}^{\text{HGDA}} = \bar{\mathbf{x}}_a^{\text{HGDA}}\mathbf{v}^T + \mathbf{X}_a^{\text{EnKF}}, \tag{6}$$

where  $\mathbf{v}$  is a vector of ones.

The flowchart of scenario b shown in Fig. 1 is similar to scenario a but with the difference that the same background mean state ( $\bar{\mathbf{x}}_b$ ) is used in both the EnKF and the VAR systems. The tunable parameter  $\alpha$  is needed as well to hybridize the analysis mean state of the EnKF and the VAR analysis. Scenario b allows computing the EnKF and the VAR solutions concurrently and also interpreted as a means of sampling data assimilation uncertainty (Houtekamer et al. 2018).

Scenario a conceptually uses the VAR to correct the analysis mean state of the EnKF, while scenario b mirrors traditional HCDA approaches and is more computationally economical. Both scenarios have been applied to numerical models with different complexities. In particular, scenario a has been applied to the NCEP operational ocean model (Penny et al. 2015; Penny 2017) while scenario b was applied to the ECMWF operational atmospheric model (Bonavita et al. 2015), and the Canadian Meteorological Centre atmospheric model (Houtekamer et al. 2018). In this study, we adopt scenario a and focus on examining how the variational optimization can improve an analysis that has already been computed by the EnKF.

In addition, it is expected that the ensemble perturbations will grow upon the unstable-neutral subspace associated with the hybrid ensemble mean state during the subsequent forecast and cycling of the DA system. Yang et al. (2012a) and Chang et al. (2014) indicated that recentering an EnKF toward a more accurate mean state improves representation of the dynamical uncertainties that are used to form the background error covariance matrix, thus better maintaining the Gaussianity of the ensemble distribution. We should emphasize that the HGDA algorithm does not “reuse” observations and thus does not violate the basic assumption that background and observation errors are uncorrelated. Since the gain covariance matrices are hybridized, the HGDA algorithm should be regarded as applying a single modified Kalman gain matrix to the innovation vector (Penny 2014). The observation information is partially used in each component DA system.

*b. HGDA with VAR update limited to the orthogonal subspace*

It is no doubt that the determination of hybrid coefficients is crucial for the performance of hybrid schemes. However, it is difficult to optimize the hybrid weighting parameter, particularly for complex systems modeling large-scale geophysical flow. This study seeks a modified HGDA algorithm to maintain the advantage of the hybrid without using a weighting parameter. With consideration for developing a maintainable DA system, we project the VAR correction to the subspace that is orthogonal to the linear subspace defined by the analysis ensemble perturbations. Thus, the variational correction to the EnKF analysis mean is only applied in directions that cannot be represented by the analysis ensemble perturbations as a basis.

To achieve this, we introduce an algorithm that we denote as the QR-HGDA. The analysis field of the VAR is appended as an additional ensemble member to the EnKF analysis ensemble (i.e., the last column vector) to form an augmented matrix  $\mathbf{A} = [\mathbf{x}_a^1, \mathbf{x}_a^2, \dots, \mathbf{x}_a^{k-1}, \mathbf{x}_a^k, \mathbf{x}_a^{\text{VAR}}]$ . All the vectors ( $\mathbf{x}$ ) are full states and the dimension of this augmented matrix is  $m \times (k + 1)$ . The QR factorization (Golub and Van Loan 2013) decomposes a matrix  $\mathbf{A}$  into a product of an orthogonal matrix  $\tilde{\mathbf{Q}}$  and an upper triangular matrix  $\tilde{\mathbf{R}}$  ( $\mathbf{A} = \mathbf{Q}\mathbf{R}$ ). By applying the QR factorization to the augmented matrix formed from the ensemble members and VAR analysis, we obtain the matrix  $\tilde{\mathbf{Q}}$  of dimension  $m \times (k + 1)$  and the matrix  $\tilde{\mathbf{R}}$  of dimension  $(k + 1) \times (k + 1)$ . All the column vectors of the  $\tilde{\mathbf{Q}}$  matrix are orthogonal to each of the column vectors. Each column vector of  $\mathbf{A}$ ,  $\mathbf{x}_a^i$  can be represented by a linear combination of the first  $i$  columns of  $\tilde{\mathbf{Q}}$  matrix. Therefore, the

last column vector in the matrix  $\tilde{\mathbf{Q}}$  is associated with the variational increment. We denote this orthogonal component as  $\mathbf{x}'_q$  and can be viewed as

$$\mathbf{x}'_q = \mathbf{K}_q (\mathbf{y}^o - \mathbf{H}\tilde{\mathbf{x}}_a^{\text{EnKF}}), \tag{7}$$

where the matrix  $\mathbf{K}_q$  is the gain matrix restricted to the subspace that is orthogonal to the analysis ensemble subspace. We note that  $\mathbf{K}_q$  can be regarded as  $\alpha\mathbf{K}_B$  and that goes back to the scenario a in Eq. (1). This study adopts the modified Gram–Schmidt algorithm (MGS, Leon et al. 2013) to extract the orthogonal component. MGS produced vectors whose inner products are closer to zero numerically than those generated by using the classical Gram–Schmidt (Leon et al. 2013).

The orthogonal component contains the variational information and is added to the EnKF analysis mean state directly to form the hybrid analysis mean in the QR-HGDA,

$$\tilde{\mathbf{x}}_a^{\text{HGDA}} = \tilde{\mathbf{x}}_a^{\text{EnKF}} + \mathbf{x}'_q. \tag{8}$$

Substituting Eqs. (2) and (7) into Eq. (8):

$$\tilde{\mathbf{x}}_a^{\text{HGDA}} = \tilde{\mathbf{x}}_b + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\tilde{\mathbf{x}}_b) + \mathbf{K}_q \{ \mathbf{y}^o - \mathbf{H}[\tilde{\mathbf{x}}_b + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\tilde{\mathbf{x}}_b)] \}, \tag{9}$$

and reformulating Eq. (9):

$$\tilde{\mathbf{x}}_a^{\text{HGDA}} = \tilde{\mathbf{x}}_b + (\mathbf{K} + \mathbf{K}_q - \mathbf{K}_q\mathbf{H}\mathbf{K})(\mathbf{y}^o - \mathbf{H}\tilde{\mathbf{x}}_b), \tag{10}$$

the hybrid gain matrix can be reformulated as

$$\hat{\mathbf{K}} = \mathbf{K} + \mathbf{K}_q - \mathbf{K}_q\mathbf{H}\mathbf{K}. \tag{11}$$

Similar to the HGDA, the hybrid gain matrix of the QR-HGDA contains a contribution of  $\mathbf{K}$  and  $\mathbf{K}_q$  with an additional term relating to any interaction between the two gain matrices. But for this formulation, all coefficients are equal to 1, with no explicit weighting parameter.

The processes of the QR-HGDA are listed as follows:

- Step1: Execute a standard EnKF for dynamical correction (the green part in Fig. 1).
- Step 2: Perform a variational analysis using the EnKF analysis mean state as the background field for climatological correction via the variational method (the red part in Fig. 1).
- Step 3: Concatenate the analysis ensemble members ( $m \times k$ ) from the EnKF and the VAR analysis field ( $m \times 1$ ) into the larger matrix  $\mathbf{A} [m \times (k + 1)]$ . Then apply the QR factorization to this augmented matrix to obtain the orthogonal component ( $\mathbf{x}'_q$ ).

Step 4: Add the orthogonal component to the EnKF analysis mean state to form a hybrid analysis mean state (the purple part in Fig. 1).

Step 5: Recenter the EnKF perturbation into the hybrid analysis mean state.

### 3. Model and experiment design

#### a. General setup

All data assimilation experiments are based on the quasigeostrophic (QG) model (Rotunno and Bao 1996). Variations of the QG model have been widely used for testing new DA methods (Hamill and Snyder 2000; Snyder et al. 2003; Corazza et al. 2003, 2007; Carrassi et al. 2008; Yang et al. 2009a,b, 2015). Our QG model configuration has a periodic channel in the zonal direction, an impermeable boundary in the meridional direction, and a rigid surface at the top and bottom. There is no terrain represented in this model. Dynamical processes include advection, diffusion, relaxation, and Ekman pumping at the bottom level. The model fields are discretized into 64 and 33 grid points in the zonal and meridional directions, respectively, with seven vertical levels. Model variables are nondimensionalized, including the potential temperature at the top and bottom levels and pseudo-potential vorticity at the five interior levels.

We use the local ensemble transform Kalman filter (LETKF; Hunt et al. 2007) as our EnKF variant. The LETKF and 3DVAR have been implemented with the QG model by Yang et al. (2009a) and Morss (1999), respectively. Observations emulate rawinsondes, containing the horizontal component of wind velocity and temperature at all levels. There are 64 observations (about 3% coverage) located at model grid points that are selected randomly at the beginning and fixed afterward. The observation locations are the same in all experiments. Rawinsonde-like observations are simulated every 12 h by adding a Gaussian distributed observation error to the true wind velocity and temperature, derived by applying a linear observation operator to the nature run.

We follow Yang et al. (2009a) to configure these DA systems with the QG model. The LETKF assimilates all observations in a local volume simultaneously to form an analysis at each grid point. The optimal dimension of the local volume depends on the ensemble size and observation density. For this QG model configuration, Yang et al. (2009a) used a horizontal localization of  $19 \times 19$  grid points with an ensemble size of 40 members. Because the ensemble size used in this study is smaller than 40, we use a localization of  $15 \times 15$  grid points for

all LETKF experiments. Although the localization is not optimized for each of the ensemble sizes, the choice of localization ensures stability with 20 members. A vertical dependent multiplicative variance inflation is applied to consider the error characteristic at each vertical level [listed in Table 2 in Yang et al. (2009a)].

To characterize the behavior of the QR-HGDA, we conduct a series of observing system simulation experiments (OSSEs). We commence experiments using the DA configuration described above, varying ensemble size (section 4). A series of sensitivity experiments are then presented to evaluate the impact of parameter changes (section 5).

#### b. Setup of the sensitivity experiments

In the first set of sensitivity experiments, we evaluate how these DA systems perform when model bias is imposed during the forecast. The model bias is generated by adjusting the vertical eddy diffusion, which impacts the Ekman pumping effect in the QG model. The QG model is forced by relaxation to a zonal mean state at all levels and an Ekman pumping at the bottom level. By decreasing the vertical eddy diffusion from 5 to  $4.75 \text{ m}^2 \text{ s}^{-1}$ , the imperfect QG model has smaller climatological variability (0.7% of the potential temperature at the bottom level) (Yang et al. 2009b). We note that observations sampled from the nature run remain unbiased.

The second set of sensitivity experiments are performed to examine how the HGDA methods respond to errors in the variational correction. For this purpose, we modify the noise applied to the synthetic observations generated from the nature run while keeping the estimated observational error covariance matrix  $\mathbf{R}$  unchanged. We evaluate the experiments where the true observation error is zero, identical to the diagonal elements in  $\mathbf{R}$ , or double these values. The purpose of these experiments is to mimic situations of overestimation, accurate estimation, and underestimation of the observation error. Inflating the observation error variance is a strategy adopted in operations (Bormann et al. 2016) to avoid overweighting correlated observations during the cost-function minimization in the VAR component. For evaluation purposes, assimilating perfect observations allows the DA to constrain the unstable dynamical error modes without introducing additional uncertainty into the innovation due to noise from random processes.

The third set of sensitivity experiments aims to understand how hybrid systems are affected by the structure of the climatological background error covariance matrix  $\mathbf{B}$ . In Morss (1999), the  $\mathbf{B}$  matrix is generated by the true 3DVAR background error with assumptions

and we refer it to as the original  $\mathbf{B}$ . We construct a new background error covariance by using the true 12-h forecast error from a long-term cycled LETKF with 40 ensemble members. Both the original  $\mathbf{B}$  and the LETKF-derived  $\mathbf{B}$  matrices are constructed in spectral coordinates, represented as

$$\mathbf{B} = \mathbf{C}^{1/2} \mathbf{V} \mathbf{C}^{1/2}. \tag{12}$$

The matrix  $\mathbf{C}$  is the horizontal background error covariance at each level and the  $7 \times 7$  matrix  $\mathbf{V}$  is the background error correlations between the vertical levels. The amplitude of  $\mathbf{B}$  can be further adjusted by multiplying an amplitude factor ( $\beta$ ) to optimize the performance of 3DVAR:

$$\tilde{\mathbf{B}} = \beta \mathbf{B}. \tag{13}$$

Following Parrish and Derber (1992), the original  $\mathbf{B}$  matrix is generated with several assumptions, including the  $\mathbf{B}$  matrix is diagonal in horizontal spectral coordinates and the  $\mathbf{B}$  matrix has separable vertical and horizontal structures and simple vertical correlations (Morss 1999), and Yang et al. (2015) suggests that a 40% reduction of the amplitude of background error covariance can optimize the performance of 3DVAR. In the sensitivity experiments, different amplitude factors and different  $\mathbf{B}$  matrices are tested to examine how the HGDA/QR-HGDA is sensitive to the choice of the  $\mathbf{B}$  matrix. Through a spectral analysis, we expect to determine the proper characterization of the static background error covariance used in HGDA.

#### 4. Results with default DA setup

We first examine the performance of the standard LETKF, while varying the ensemble size, by evaluating the root-mean-square error (RMSE) of potential temperature at the bottom level. As shown in Fig. 2, when the ensemble size ( $k$ ) is equal to 5, the standard LETKF immediately experiences filter divergence and fails to track the trajectory of nature run. Increasing the ensemble size to 6, the RMSE is smaller at first but eventually diverges as well. Generally, it is expected that a larger ensemble size will better represent the growing modes of errors. However, Fig. 2 shows that the LETKF is stable with  $k = 7$  but the RMSE jumps abruptly when increasing to an ensemble size  $k = 8$ . Nevertheless, the LETKF with  $k = 8$  converges to a smaller RMSE than the one using  $k = 7$  after day 84. It is possible that the experiment with ensemble size  $k = 7$  was able to capture the dominant error mode between days 45–50 by chance, while the ensemble size  $k = 8$  was not. It is clear

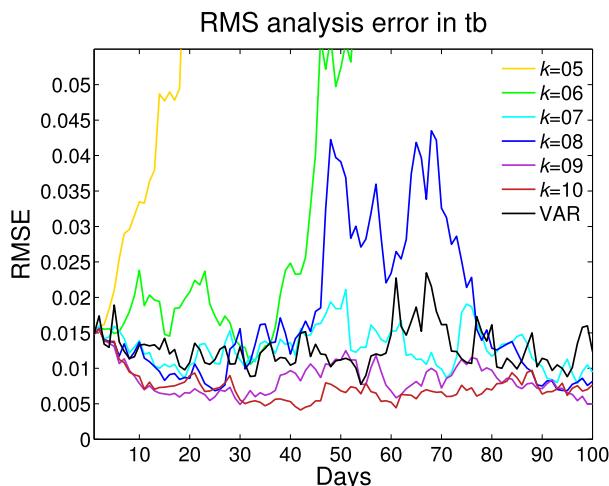


FIG. 2. Time series of the RMSE of VAR and LETKF with different ensemble sizes ( $k$ ) in a perfect model assumption.

that the LETKF is divergent for ensemble size  $k \leq 6$  and stable for ensemble size  $k \geq 9$ . Thus we consider the cases of  $k = 7$  and  $k = 8$  as a transition zone of stability. Compared to the LETKF, the 3DVAR provides a stable result but the overall accuracy (mean RMSE = 0.012254) is worse than LETKF with  $k = 9$  members (RMSE = 0.008437).

Figure 3 illustrates the performance of the HGDA system with  $\alpha = 0.5$ . The HGDA recovers stability with  $k = 5$  and  $k = 6$  compared with the standard LETKF diverging at these ensemble sizes. Moreover, although the RMSE increases during some particular periods, the HGDA avoids filter divergence even with an ensemble size as small as  $k = 3$ . These results agree qualitatively with the findings for the Lorenz model by Penny (2014). However, we note that the performance of the HGDA demonstrated here is sensitive to the static background error covariance matrix. This sensitivity will be discussed further in section 5c.

The EnKF requires an accurate ensemble mean and a representative set of perturbations that have converged toward the unstable-neutral error subspace. For this reason, the EnKF tends to require a longer spinup as compared with 3DVAR or 4DVAR (Kalnay and Yang 2010; Houtekamer and Zhang 2016). Comparing the LETKF (purple line in Fig. 2) and HGDA (purple line in Fig. 3) with ensemble size  $k = 9$ , the HGDA exhibits an accelerated spinup period (e.g., HGDA RMSE = 0.0073 versus LETKF RMSE = 0.0088 at day 8) due to the improved accuracy ensemble mean state (Yang et al. 2012b).

To illustrate the impact of using different ensemble sizes and different DA algorithms, the same background ensemble is used in the following discussion.

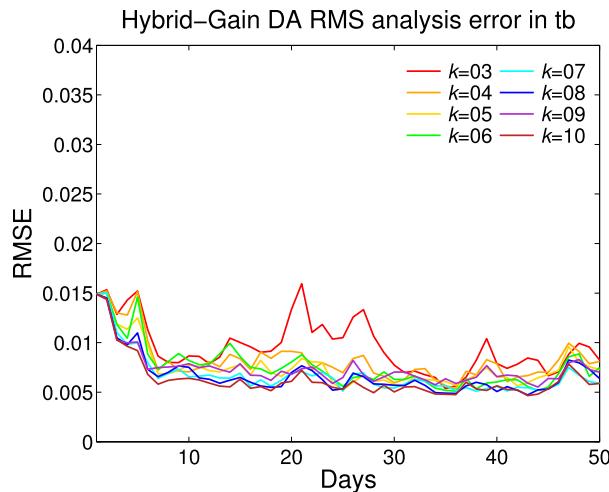


FIG. 3. Times series of the RMSE of HGDA ( $\alpha = 0.5$ ) with different ensemble sizes ( $k$ ) in a perfect model assumption.

The background ensemble is generated by performing the standard LETKF with 40-members for 20 days. The expected background error is represented by the difference between the mean of the 40-member background ensemble and the nature run. Based on this 40-member ensemble, we randomly select 6 members (left panel of Fig. 4) and 30 members (right panel of Fig. 4), respectively, and then apply the standard LETKF and HGDA one time without cycling. The observation locations are marked as green dots in Fig. 4a. We first focus on the area around the point ( $x = 30, y = 5$ ), where there is a negative error in the background field. In Fig. 4a, the positive increment in that region indicates that LETKF is able to correct this background error. However, with the ensemble size  $k = 6$ , the correction only applies to half of the spatial extent of the error in that area. When the ensemble size is increased, the correction in Fig. 4d covers the entire spatial extent corresponding to the negative error structure. The extensively wrong correction around the point ( $x = 25, y = 12$ ) in Fig. 4a also vanishes with an increased ensemble size. Figures 4b and 4e describe the VAR correction based on the corresponding LETKF analysis mean state. Figures 4b and 4e exhibit similar structure due to the use of the same static background error covariance matrix, and tend to have an isotropic-like structure, compared to the LETKF analysis increments. The VAR corrections can also be useful, such as the area near the point ( $x = 27, y = 17$ ).

Figure 4c shows the orthogonal correction of the QR-HGDA. Without setting a hybrid coefficient, the QR-HGDA achieves a similar amount of correction. There are still slight differences between the VAR correction of the HGDA (Fig. 4b) and the extracted orthogonal

component (Fig. 4c). Figure 5 highlights these differences by subtracting the analysis increment of VAR from the extracted orthogonal component. These differences become more evident as the ensemble size increases from  $k = 6$  to  $k = 30$  (Fig. 5a versus Fig. 5b). As the ensemble size increases, the error space can be represented more completely, while the dimension of the orthogonal space is reduced and representative of less dominant error modes.

Figure 6 shows the performance of the cycled HGDA system with the ensemble size  $k = 6$  at day 20. We focus on the area between ( $x = 25$  to  $x = 45, y = 25$  to  $y = 33$ ), where there is a large positive background error structure. The LETKF, due to its limited ensemble size, only corrects the left half error structure (Fig. 6a) while the 2nd step update in the HGDA corrects the right half part (Fig. 6b). The total HGDA analysis increment is given in Fig. 6c. A comparison between the VAR increment and the orthogonal correction of the QR-HGDA shows a similar structure, indicating a similar ability to correct the background error structure.

While there is little error in the background around the point ( $x = 50, y = 30$ ), both the VAR increment (Fig. 6b) and the orthogonal correction (Fig. 6d) reveal an erroneous adjustment. As shown in Fig. 6c, the inaccurate corrections made by the VAR increment are damped by the weighted combination with the LETKF analysis. However, the corresponding error remains in the QR-HGDA (Fig. 6e). Since those inaccurate error structures do not occur in the ensemble subspace, applying them in the space orthogonal to the ensemble simply adds noise to the cycled HGDA system. The QR-HGDA directly reflects the correction provided by the VAR and thus fully inherits its inaccuracy. Therefore, improper estimation of the static error covariance matrix may degrade the accuracy of QR-HGDA. This provides an explanation for why the QR-HGDA always has a higher RMSE than the HGDA in all of the experiments presented so far (Table 1). We will next utilize this property to diagnose the quality of the static background error covariance within the hybrid.

Table 1 presents a 50-day average RMSE for the LETKF, the HGDA, and the QR-HGDA with ensemble sizes ranging from  $k = 3$  to  $k = 20$ . Note that the RMSEs in all tables are multiplied by a factor 100. For the smallest ensemble sizes, the HGDA produces the lowest RMSE, followed by the QR-HGDA. Both hybrid methods prevent filter divergence when the ensemble size is small ( $k < 9$ ) and produce lower RMSE than the standard LETKF. As expected, the advantage of the QR-HGDA over the LETKF to eliminate filter divergence vanishes as the ensemble size increases ( $k = 9$ ) and LETKF becomes more stable. When the ensemble

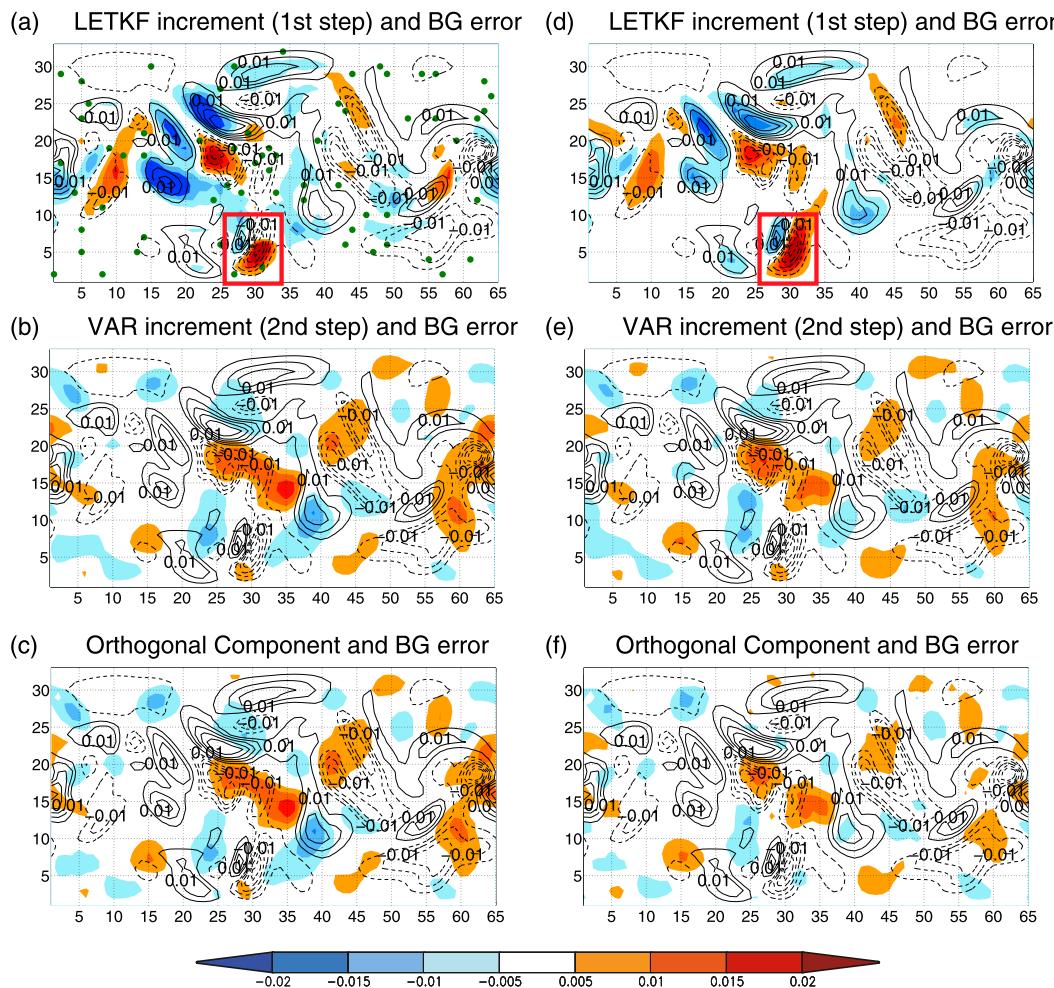


FIG. 4. Snapshots of the potential temperature at the bottom level of the same background field experiment at day 20 with different ensemble sizes ( $k = 6$  and  $k = 30$ ). Background error (contour, the dashed and solid lines represent the negative and positive value, respectively) and analysis increment (shade) of (a) LETKF with  $k = 6$ , (b) VAR (the second step update in HGDA; its background field comes from the analysis mean state of EnKF), and (c) orthogonal component of QR-HGDA. (d)–(f) As in (a)–(c), but using ensemble size  $k = 30$ . The observation locations are marked as green dots in (a).

size increases to  $k = 20$ , the background error covariance generated by LETKF adequately represents the growing error modes for the purpose of maintaining filter stability. With an ensemble size of  $k = 20$ , the variational corrections using a static background error covariance matrix are overestimated and degrade the performance of both hybrids. The QR-HGDA is more sensitive to these inaccuracies compared to the HGDA. For example, the RMSE for the QR-HGDA with  $k = 20$  is even larger than the HGDA with  $k = 5$ .

Since the purpose of the QR-HGDA is to avoid using the hybrid weight, it is essential to evaluate the QR-HGDA compared to HGDA with a varying hybrid weighting value. Figure 7 shows the RMSE of the HGDA as a function of the combination weight ( $\alpha$ ) and

with different ensemble sizes versus the RMSE of the QR-HGDA. The lowest RMSE for ensemble sizes  $k = 3$ ,  $k = 6$ , and  $k = 10$  occur with  $\alpha = 0.6$ ,  $\alpha = 0.4$ , and  $\alpha = 0.2$ , respectively. As expected, the optimal weighting parameter depends on the ensemble size, and less correction from the VAR is needed as the ensemble size increases. As demonstrated with Figs. 4c and 4f, the orthogonal component also becomes smaller as more ensemble members are used. With  $k = 10$ , the HGDA always leads to smaller RMSE than the QR-HGDA regardless of the weighting parameter.

At these three small ensemble sizes, the HGDA becomes more sensitive to the tuning parameter with RMSE ranging from 0.006 to 0.22, while the QR-HGDA is relatively robust with RMSE ranging from 0.0082 to 0.0119.

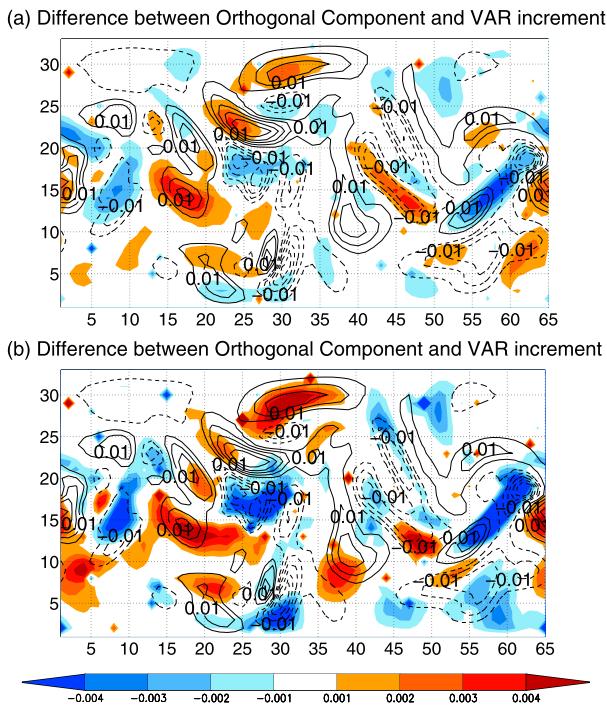


FIG. 5. Snapshots of the potential temperature at the bottom level of the same background field experiment at day 20. Background error (contour, the dashed and solid lines represent the negative and positive value, respectively) is the same as Fig. 4 and the difference (shade) between analysis increment of VAR and the orthogonal component of QR-HGDA with ensemble size (a)  $k = 6$  and (b)  $k = 30$ .

When the ensemble size is small ( $k = 6$ ) and uses a small weight value (i.e., more impact from the LETKF, e.g.,  $\alpha = 0.1$ ), the QR-HGDA provides a slight improvement over the HGDA. The HGDA becomes unstable if the ensemble size is small (e.g.,  $k = 3$ ) and  $\alpha$  is small. There is a benefit of QR-HGDA over HGDA when  $\alpha \leq 0.5$ .

## 5. Results with sensitivity experiments

Errors in the DA process include systematic sampling error of the ensemble, imbalances caused by localization, and improper estimation of the error covariance matrices. Three sensitivity experiments (setup for the sensitivity experiment is described in section 3b) are conducted to understand how the HGDA and the QR-HGDA respond to these issues by including model bias, misestimation of observation error covariance, and misestimation of background error covariance.

### a. Sensitivity to model bias

We have demonstrated under a perfect model framework that both HGDA and QR-HGDA prevent the filter divergence that occurs with LETKF when using

small ensemble sizes. We now examine how the DA schemes perform in the presence of model bias. Figure 8 shows the performance of the LETKF systems implemented with an imperfect QG model. Compared to the results using a perfect model (Fig. 2), the variations in RMSE are larger. For  $k = 8$ , the RMSE grows rapidly after day 30 and the filter diverges. Similarly to the perfect model, nine ensemble members are required to prevent filter divergence with the imperfect QG model, but there are additionally two high peaks in RMSE between days 42 to 63.

Table 2 displays the 50-day average performance of the LETKF, HGDA, and QR-HGDA using an imperfect QG model with ensemble sizes ranging from  $k = 3$  to  $k = 20$ . The weighting parameter used in the HGDA is  $\alpha = 0.5$ . Unlike the results for the perfect QG model, where the HGDA always produces smaller RMSE than the QR-HGDA for all ensemble sizes, the QR-HGDA with the imperfect QG model has smaller RMSE and bias than the HGDA for the small ensemble sizes  $k = 3$  and  $k = 4$ . This indicates that with a larger ensemble size (e.g.,  $k = 20$ ), the HGDA still reduces RMSE compared to the LETKF, indicating that the climatological background error covariance still provides value. However, similar to what has been discussed in relation to Table 1, the QR-HGDA has worse performance than the LETKF for  $k = 20$ .

### b. Sensitivity to observation error estimation

The quality of the observations and whether the observation error is adequately represented can affect the DA performance. Penny (2017) discussed the impact of observation noise on the stability of the filter in relation to the Lyapunov exponents of the DA systems. Penny (2017) found that a 3DVAR system with a leading Lyapunov exponent that is negative but small in magnitude eventually diverges from the truth due to noise in the observations. It indicates that observation noise can impact the stability of the DA systems. Under the perfect model setup, we examine how the presence of noise in the observations and the improper estimation of this noise in the observation error covariance matrix can affect the stability and accuracy of the DA systems. These sensitivity experiments are performed with three different sets of observations, but the same observation error covariance ( $\mathbf{R}$ ) is prescribed. Observations are generated as 1) error free (i.e., zero noise, with observations directly sampled from the nature run), 2) with accurate error variance corresponding to  $\mathbf{R}$ , and 3) with error variance double that represented by  $\mathbf{R}$ . These experiments will be used to indicate situations of overestimation, accurate, and underestimation of  $\mathbf{R}$ , respectively.

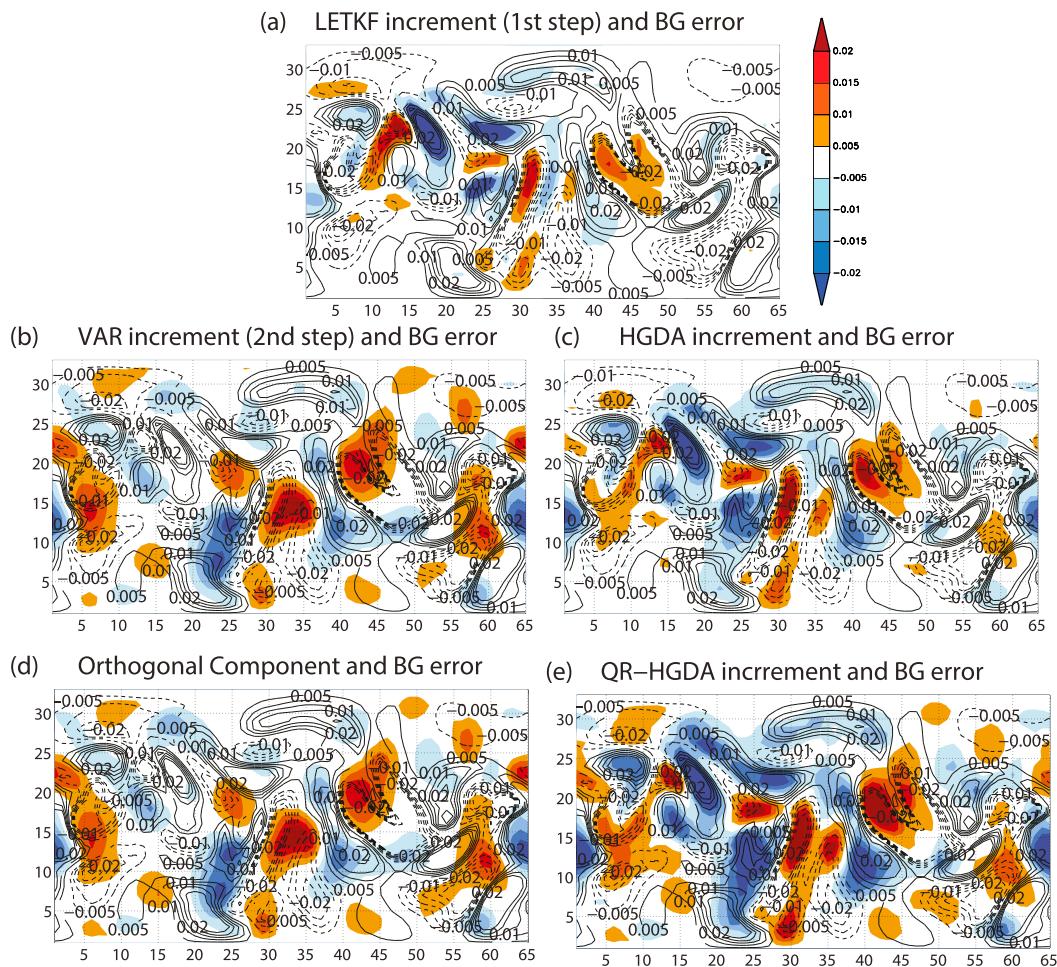


FIG. 6. Snapshots of the potential temperature at the bottom level for cycling DA experiment with 6 ensemble members at day 20. Background error (contour) and analysis increment (shade) of (a) LETKF, (b) VAR (the second step update in HGDA), (c) HGDA (total correction), (d) orthogonal component, and (e) QR-HGDA (total correction).

As a baseline, with zero observation noise, the RMSE of all DA systems is reduced dramatically (e.g., Table 3), confirming that observation noise has a large impact on the accuracy of the DA systems. When observation noise

is not present, LETKF can confine the error growth rate even with a small ensemble size. For example, without noise the RMSE with six members is only 0.006595 compared to the very large RMSE with the 6-member LETKF

TABLE 1. 50-day average RMSE for LETKF, HGDA, QR-HGDA, and local QR-HGDA with different ensemble sizes ( $k$ ) using a perfect model assumption. Note the RMSEs are multiplied by a factor 100.

50-day avg RMSE ( $\times 100$ )	LETKF	HGDA	Global QR-HGDA	Local QR-HGDA
$k = 3$	33.7622	0.9591	1.0329	N/A
$k = 4$	28.7014	0.8248	0.9876	2.3357
$k = 5$	19.2092	0.7563	0.9375	2.0828
$k = 6$	2.2310	0.7649	0.9201	1.6727
$k = 7$	1.3271	0.6750	0.8615	1.1881
$k = 8$	1.4456	0.6767	0.8848	1.2120
$k = 9$	0.8437	0.6562	0.8461	1.0229
$k = 10$	0.7667	0.6345	0.8205	0.8783
$k = 15$	0.6549	0.6147	0.8072	0.7935
$k = 20$	0.5436	0.5934	0.7738	0.6667

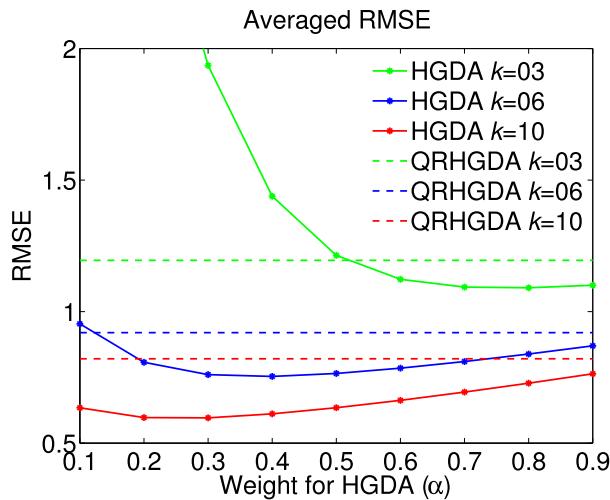


FIG. 7. RMSE for HGDA (solid lines) and QR-HGDA (dash lines) with ensemble size  $k = 3$  (green line),  $k = 6$  (blue line), and  $k = 10$  (red line) as a function of combination weight ( $\alpha$ ) implemented with a perfect QG model. Note that the RMSEs are multiplied by a factor of 100.

with observation noise. When observation noise is present, similar accuracy can only be achieved by using at least 15 ensemble members (Table 1). The HGDA further improves the accuracy compared to the LETKF. Now, however, the QR-HGDA produces the most accurate analyses out of all the DA methods examined. This benefit is most evident at small ensemble sizes (Fig. 9).

Using the same background ensemble as in Fig. 4, Fig. 10 shows the background error and analysis increments for the 6-member LETKF, the second-step variational update of the HGDA, and the orthogonal component of the QR-HGDA, while varying the observation noise. When the observational noise increases from zero (implying an overestimation of  $\mathbf{R}$  in the data assimilation) to the correct value (accurate estimation) to the doubled value (underestimation of  $\mathbf{R}$ ), the increment structure of the LETKF update does not change much and is limited to certain regions. The amplitude of the increments increases as the observation noise increases (Fig. 10a versus Fig. 10g) because this noise only appears in the innovation term (recall the same  $\mathbf{R}$  matrix is used in all experiments). When the observation noise is doubled, the prescribed  $\mathbf{R}$  matrix underestimates the inaccuracy of the innovation term. The underestimation of the gain matrix can lead to the underestimation of the amplitude of the increment, and vice versa. However, the increment structure of both the VAR and the orthogonal component emerges in more areas, introducing more detrimental adjustments caused by the observation noise (e.g.,  $x = 53$ ,  $y = 25$ ). This suggests that the variational correction is sensitive to the observation

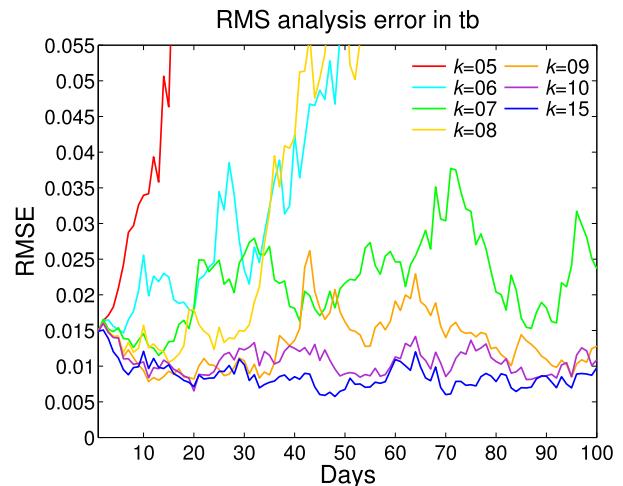


FIG. 8. Time series of the RMSE of LETKF with different ensemble sizes ( $k$ ) in an imperfect (biased) QG model. It can be compared with the perfect QG model experiment shown in Fig. 2.

noise, and inaccurate corrections are further emphasized by the QG-HGDA. The LETKF can ignore noise in directions that do not align with the growing error modes represented by the ensemble perturbations, which makes it less sensitive to the statistically isometric observational noise.

To further elucidate this phenomenon, we also perform the assimilation of perfect observations for either only the LETKF or only the VAR component of the HGDA and QR-HGDA (Table 3). The average RMSE of assimilating the perfect observations in the 1st step and the appropriate observation noise in the 2nd step is larger than the vice versa case (zero noise in LETKF and default noise in VAR in Table 3). These results indicate that a significant portion of the inaccurate corrections of the VAR component can be attributed to sensitivity to observational noise. Penny (2017) demonstrated with the Lorenz 96 model that to avoid filter divergence, the leading Lyapunov exponent (LLE) of the DA system must be sufficiently negative to avoid the destabilizing impacts of observational noise. In other words, not only must the leading Lyapunov exponent (LLE) of the DA system be negative, but attention must also be paid to stable modes that have the LLE close to zero. A small magnitude LLE renders the DA system more susceptible to noise in the observations, increasing the probability for the DA system to become unstable. The HGDA can recover stability by driving the LLE sufficiently negative to be insensitive to small levels of observational noise.

### c. Sensitivity to background error estimation

Experiments in the previous subsection revealed that the QR-HGDA is slightly less accurate than the

TABLE 2. 50-day average RMSE and the averaged absolute bias (in parentheses) for LETKF, HGDA, and QR-HGDA with different ensemble sizes ( $k$ ) using an imperfect model assumption. Note the RMSEs are multiplied by a factor 100.

50-day avg RMSE ( $\times 100$ )	LETKF-bias	HGDA-bias	QRHGDA-bias
$k = 3$	33.6292 (0.0220)	1.2137 (0.0009)	1.1950 (0.0008)
$k = 4$	23.2106 (0.0316)	1.1113 (0.0011)	1.0936 (0.0010)
$k = 5$	16.6046 (0.0156)	0.9482 (0.0007)	1.0530 (0.0007)
$k = 6$	2.9232 (0.0087)	0.9226 (0.0008)	1.0190 (0.0009)
$k = 7$	1.8970 (0.0020)	0.8283 (0.0005)	0.9571 (0.0005)
$k = 8$	2.5857 (0.0041)	0.8353 (0.0007)	0.9695 (0.0008)
$k = 9$	1.2368 (0.0008)	0.8003 (0.0007)	0.9399 (0.0008)
$k = 10$	1.1039 (0.0014)	0.7663 (0.0008)	0.8962 (0.0008)
$k = 15$	0.9019 (0.0007)	0.7372 (0.0005)	0.8840 (0.0005)
$k = 20$	0.7742 (0.0007)	0.7031 (0.0005)	0.8617 (0.0007)

standard HGDA unless the observations are very accurate or the observation error variance is overestimated (e.g., Table 3). An alternative view would be to consider that the ratio of the background error variance in  $\mathbf{B}$  to the observation error variance in  $\mathbf{R}$  is underestimated within the Kalman gain. Thus in attempting to tune the configuration of the QR-HGDA, it may be more practical to focus on improving the estimation of the static  $\mathbf{B}$  matrix used for the hybrid assimilation with the default  $\mathbf{R}$  matrix in the perfect model. To investigate the role of the static  $\mathbf{B}$  matrix in the HGDA and the QR-HGDA, we construct a new static background error covariance matrix from cycled LETKF forecasts (LETKF-derived  $\mathbf{B}$ ). We note that both the original  $\mathbf{B}$  and the LETKF-derived  $\mathbf{B}$  matrices are constructed with the NMC method.

To understand the characterization of the  $\mathbf{B}$  matrices, we calculate the horizontal error variance in the  $\mathbf{B}$  matrix as a function of approximated global wavenumber for the temperature at the bottom model level (Fig. 11). As noted in Morss (1999), the global wavenumber is not well defined in the channel model; however, it indicates the error variance structure at different spatial scales. The original horizontal background error structure (Fig. 11a) has a larger variance from large to middle scales (wavenumber smaller than 20) and very little variance at small scales. In comparison, the error variances from the LETKF-derived  $\mathbf{B}$  (Fig. 11b) at smaller scales are much larger. One cause of this difference is that the background error used to construct the original  $\mathbf{B}$  matrix contains both growing and nongrowing modes while the background error used for the LETKF-derived  $\mathbf{B}$  is flow-dependent and dominated by fast growing errors. Also, large error variance at small scales can reflect noise introduced by procedures such as localization (Yang et al. 2009b).

The different structure of the LETKF-derived  $\mathbf{B}$  matrix from the original  $\mathbf{B}$  matrix will lead to different performance when they are used by a variational DA

method. The amplitude of the static  $\mathbf{B}$  matrix can also affect the performance of the VAR and thus affect the hybrids. Therefore, we investigate the sensitivity of the HGDA and the QR-HGDA to both formulations of the static  $\mathbf{B}$  matrix with varying amplitudes. For the stand-alone VAR system, 0.4 is the optimal amplitude factor [ $\beta$  in the Eq. (13)] for both the original  $\mathbf{B}$  matrix (the averaged RMSE is 0.012254) and the LETKF-derived  $\mathbf{B}$  matrix (the averaged RMSE is 0.012712). The same optimal factor indicates that the background error variance is overestimated from the assumptions applied for constructing the climatological  $\mathbf{B}$  matrix (Morss 1999). Figure 12 presents the time-averaged RMSE of the HGDA and the QR-HGDA with different background error covariances as a function of the amplitude factor. In both the  $k = 6$  (Fig. 12a) and  $k = 10$  (Fig. 12b) cases, the QR-HGDA is more accurate than the HGDA when the LETKF-derived  $\mathbf{B}$  matrix is used (red lines in Fig. 12) by the VAR component. Increasing the amplitude of the LETKF-derived  $\mathbf{B}$  matrix leads to improved performance in both the HGDA and QR-HGDA, even when using larger ensemble sizes. Also, the difference between the HGDA and QR-HGDA is reduced with a larger ensemble size.

TABLE 3. 50-day average RMSE for LETKF, HGDA, and QR-HGDA with 6 members with different observation noise scenarios. Using different observation noise scenario is not applicable to the stand-alone LETKF. Note the RMSEs are multiplied by a factor 100.

50-day avg RMSE ( $\times 100$ )	LETKF	HGDA	QR-HGDA
Zero noise	0.6595	0.3717	0.3122
Default noise	2.2310	0.7649	0.9201
Double noise	2.6312	1.2961	1.7248
Zero noise in LETKF and default noise in VAR	N/A	0.6784	0.8769
Default noise in LETKF and zero noise in VAR	N/A	0.4778	0.4184

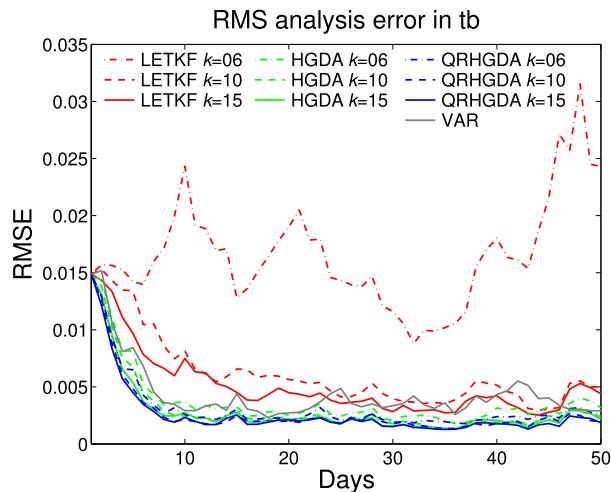


FIG. 9. Time series of the RMSE for 3DVAR (gray), LETKF (red), HGDA (green,  $\alpha = 0.5$ ), and QR-HGDA (blue) with ensemble sizes  $k = 6$  (dashed-dotted curves),  $k = 10$  (dashed curves), and  $k = 15$  (solid curves) in a perfect model and perfect observation assumption.

However, as shown in Fig. 12b, increasing the amplitude of the original  $\mathbf{B}$  matrix degrades the hybrids and results in poorer performance compared to a stable LETKF ( $k = 10$ ). This is attributed to the quick degradation at larger scales, given that the original  $\mathbf{B}$  matrix has larger power spectrum at these scales. This suggests that neither the HGDA nor the QR-HGDA gains much benefit from the large-scale information provided by the original  $\mathbf{B}$  matrix. For the LETKF, the background error and corrections are related to dynamical instabilities and to the density of the observation network used in this study (Fig. 4a). As a result, LETKF mainly corrects the structures with wavenumber smaller than 15 (Yang et al. 2009a, 2015), constraining errors at very large scales. Since the 2nd step update of HGDA is applied to the LETKF analysis mean, it is unnecessary to correct the larger scale again in the VAR when the LETKF uses a sufficient ensemble size. However, as the ensemble size is reduced to the point that it is insufficient to maintain stability, the LETKF analysis gains more value from the original  $\mathbf{B}$  matrix's representation of the error modes at large scales, even if they are poorly represented.

We perform a spectral analysis to further identify changes due to the different background error covariance estimates applied during each update step of the QR-HGDA. Figure 13 shows the time-averaged power spectra for each update step in the QR-HGDA using the original  $\mathbf{B}$  or the LETKF-derived  $\mathbf{B}$  in the variational correction. For the QR-HGDA (HGDA as well), the corrections are dominated by the LETKF (first step

update) but their characterizations are different when using different background error covariance structures in the second step update during the cycling run. In general, the corrections provided by both steps are larger with the HGDA using the original  $\mathbf{B}$  matrix than the one with the LETKF-derived  $\mathbf{B}$  matrix. For the first step update, the spectrum power of the increment with the HGDA using the original  $\mathbf{B}$  matrix has much higher variance at large scales while the power distribution with the LETKF-derived  $\mathbf{B}$  matrix has a peak between wavenumbers 10 to 20. Compared with the power spectrum distribution of the standard LETKF with 40 members (Fig. 13b), the high variance at large scales is attributed to the original  $\mathbf{B}$  matrix as a consequence of the overcorrection at these scales. For the second step update, the slope associated with the LETKF-derived  $\mathbf{B}$  matrix is smaller than that associated with the original  $\mathbf{B}$  matrix at large scales. This demonstrates that these two  $\mathbf{B}$  matrices resulted in very different bases for correcting the dominant errors in the LETKF analysis mean.

The findings in this experiment with the LETKF-derived  $\mathbf{B}$  matrix suggest that deriving a proper static  $\mathbf{B}$  matrix, which can compensate the ensemble-sampled flow-dependent  $\mathbf{B}$  matrix and help to better describe the uncertainty in the ensemble mean state, can enhance the value of hybrid methods. A well-tuned  $\mathbf{B}$  matrix for the stand-alone variational systems is not an excellent candidate for the hybrids since the original well-tuned  $\mathbf{B}$  matrix represents the uncertainties for a climatological background error characterization rather than an analysis mean state with dynamical corrections in HGDA. With the LETKF-derived  $\mathbf{B}$  matrix, both the HGDA and QR-HGDA produce smaller RMSE than the standard LETKF (shown in Fig. 12) and the QR-HGDA outperforms the HGDA, especially with  $k = 6$ . However, we note that, in addition to the structure of the  $\mathbf{B}$  matrix, there is another parameter, the amplitude of the  $\mathbf{B}$  matrix, that may affect the performance of the hybrids. Therefore, in Fig. 14, we calculate the time-averaged RMSE difference between the hybrids (QR-HGDA minus HGDA) by varying the combination weighting ( $\alpha$ ), which is used in the HGDA, and the amplitude factor ( $\beta$ ), which applied to adjust the amplitude of the static LETKF-derived  $\mathbf{B}$  matrix. In each row element (e.g.,  $\beta = 1.0$ ) of Fig. 14, QR-HGDA has only one solution but the outcome of HGDA depends on the choice of  $\alpha$ . We note that the results in column  $\alpha = 0.5$  are the same as the hybrid methods using the LETKF-derived  $\mathbf{B}$  matrix (red lines) shown in Fig. 12b. As shown in Fig. 14, in general, the QR-HGDA produces smaller or similar RMSE values to those of the

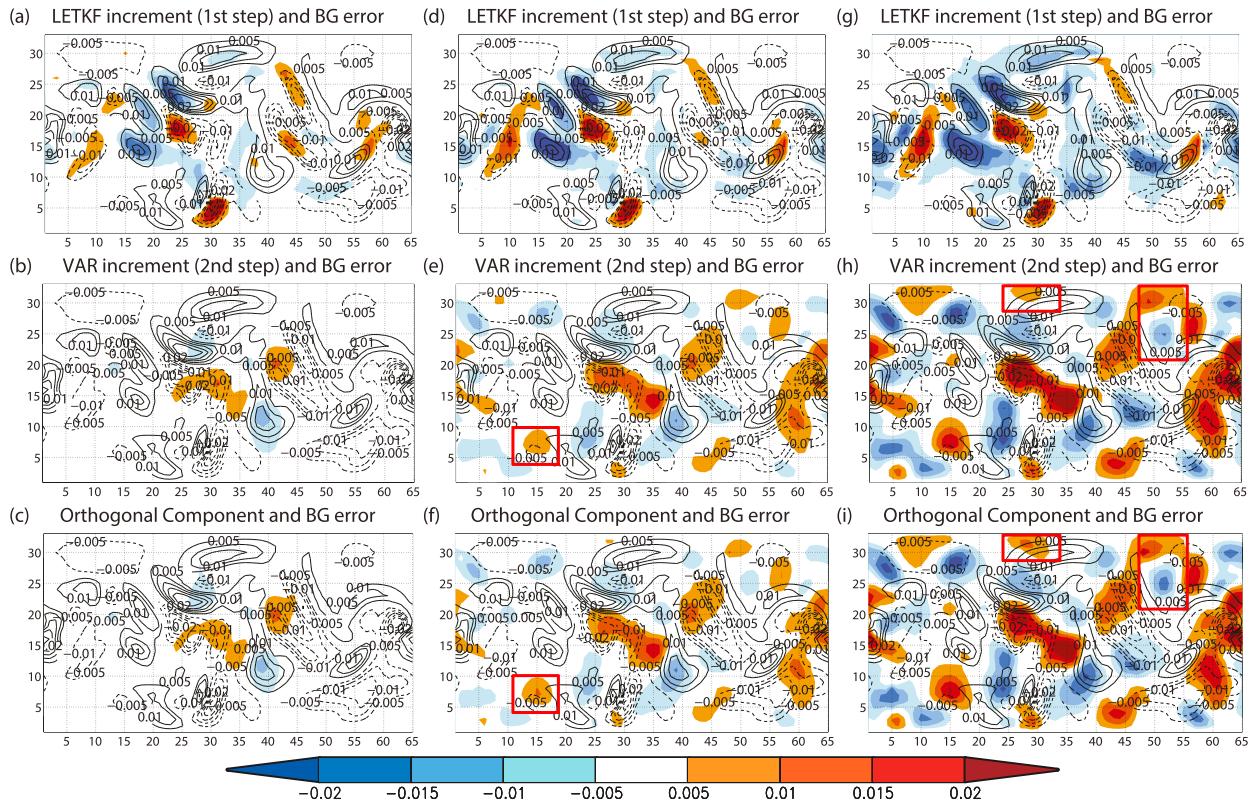


FIG. 10. Snapshots of the potential temperature at the bottom level for cycling DA experiment at day 20 with 6 ensemble members. Background error (contour) and analysis increment (shade) of (a) LETKF, (b) VAR (the second step update in HGDA), and (c) orthogonal component of QR-HGDA with observation noise equal to zero; (d)–(f) as in (a)–(c), but with observation noise equal to 1 (default value); and (g)–(i) as in (a)–(c), but using the double observation noise.

HGDA with all combination weightings  $\alpha$ . A parameter like  $\beta$  is usually determined when the static  $\mathbf{B}$  matrix is optimized for the 3DVAR system and would not be tuned for the hybrid method.

The LETKF-derived  $\mathbf{B}$  may not yet be the optimal choice for the HGDA/QR-HGDA; the main purpose of this experiment is to examine how a  $\mathbf{B}$  matrix with a different characterization in comparison to the

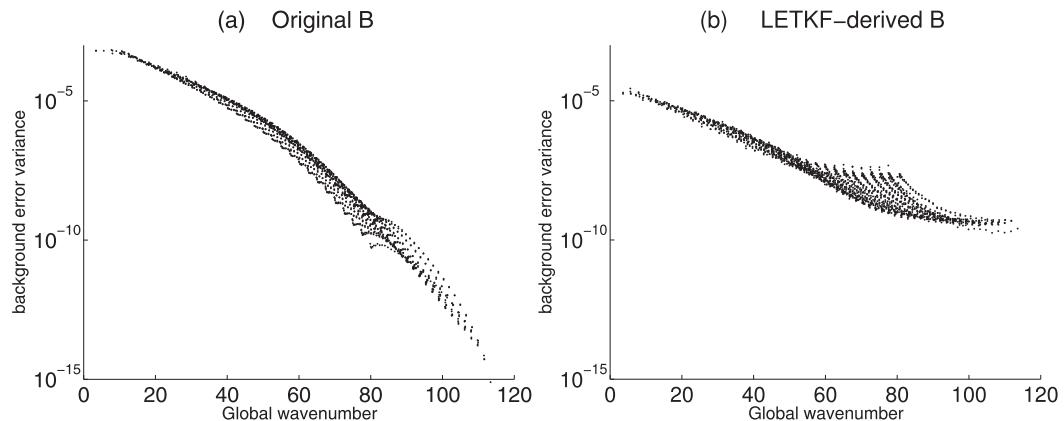


FIG. 11. Horizontal background error variances of (a) the original  $\mathbf{B}$  matrix and (b) the LETKF-derived  $\mathbf{B}$  matrix as a function of approximate global wavenumber. The global wavenumber is defined by Morss (1999) as:  $[(2.5 \times k)^2 + (5.2 \times 0.5 \times l)^2]^{1/2}$ . The  $k$  and  $l$  represent the zonal wavenumber and the meridional half-wavenumber, respectively. The factors of 2.5 and 5.2 are applied to scale the zonal and meridional extent of the QG model to the real world.

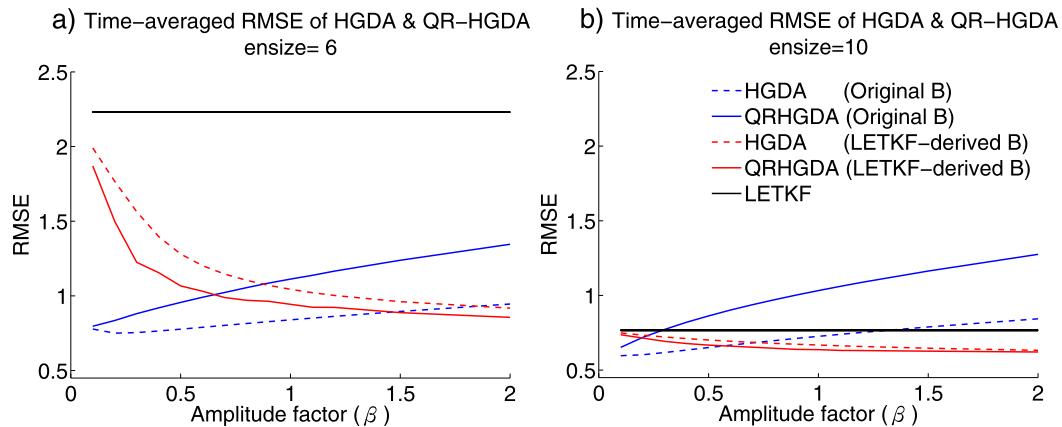


FIG. 12. The time averaged RMSE of LETKF (black solid line), HGDA (dashed line), and QR-HGDA (solid line) with ensemble size (a)  $k = 6$  and (b)  $k = 10$  as a function of different amplitude factor ( $\beta$ ) used in the original (blue lines) and the LETKF-derived (red lines) background error covariance matrix. The amplitude factor is a constant for amplifying the variance of the  $\mathbf{B}$  matrices. Note that the RMSEs are multiplied by a factor of 100.

original one can modify the performance of the hybrids. The LETKF-derived  $\mathbf{B}$  produces more stable performance for both hybrids and enhances the QR-HGDA (with no dependency on an additional hybrid weighting coefficient) compared to the HGDA. Thus, it may be more beneficial to shift the focus from the tuning of a single hybrid weighting parameter to focusing on improved estimation of the full climatological  $\mathbf{B}$  matrix. Further, we believe it is informative to examine which characteristics of the climatological  $\mathbf{B}$  matrix permit the hybrid methods to achieve their best performance.

## 6. Global orthogonality versus local orthogonality

The main mechanism of the QR-HGDA is to limit the variational correction to the EnKF ensemble mean to

the subspace orthogonal to the linear subspace defined by the analysis ensemble perturbations. Penny (2014) applied hybrid variational corrections to LETKF both locally and globally and found that the global corrections generally produced more accurate results. Thus, the orthogonalization in the experiments discussed above was conducted globally. This global orthogonality has some advantages for constraining the large-scale bias and errors using the variational analysis. Further, the information carried in the global variational correction, such as dynamical balance, would not be distorted.

Given that the forecast errors with the LETKF system are dominated by local instabilities, conducting the orthogonalization locally in QR-HGDA may emphasize the local relationship between the errors in the analysis ensemble mean from LETKF and the variational correction. In this section, we compare the performance

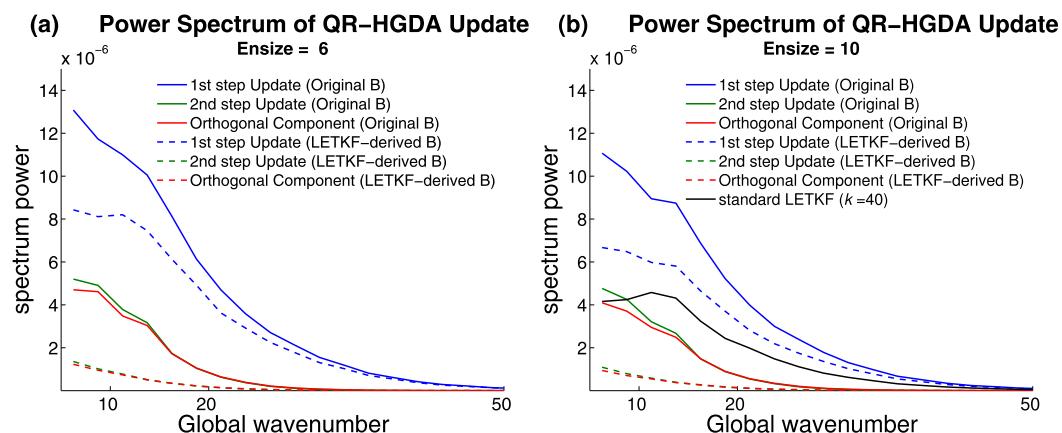


FIG. 13. Time-averaged power spectra of the increment at each update steps of HGDA and the orthogonal component of QR-HGDA with ensemble size (a)  $k = 6$  and (b)  $k = 10$ , using different  $\mathbf{B}$  matrix in the VAR system. The black line shows the time-averaged power spectra of standard LETKF increment with 40 ensemble members.

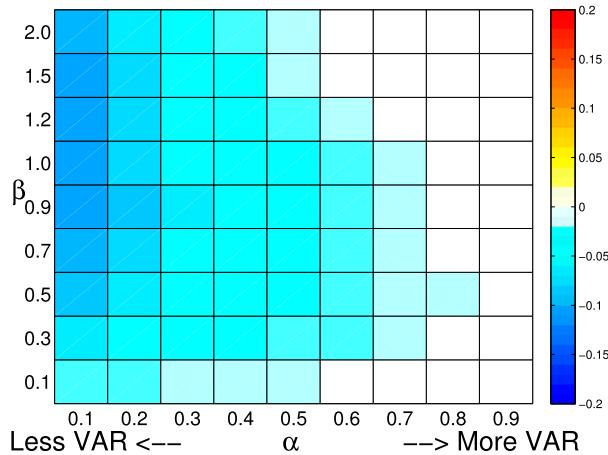


FIG. 14. RMSE difference between HGDA and QR-HGDA (the QR-HGDA minus the HGDA) using 10 ensemble members with the combination weight of HGDA varying from  $\alpha = 0.1$  to  $\alpha = 0.9$ , and the amplitude of LETKF-derived  $\mathbf{B}$  matrix from  $\beta = 0.1$  to  $\beta = 2.0$ . Negative (blue) values indicate that the QR-HGDA is more accurate than the HGDA. Note that the RMSE differences are multiplied by a factor of 100.

of QR-HGDA with global and local orthogonalization. The local orthogonalization is done by calculating the orthogonality at each local patch with the identical local radius used by LETKF, and the local orthogonal component is used to correct the corresponding grid point. Therefore, the following result also demonstrates the possibility to incorporate the variational correction based on local orthogonality, which may have a computational advantage due to potential opportunities for parallelization when extending the QR-HGDA to higher dimensions.

With the original  $\mathbf{B}$  matrix, the averaged RMSE of QR-HGDA with local orthogonality (denoted as local QR-HGDA) is shown in Table 1 and it can be compared with the result of QR-HGDA with global orthogonality. The local QR-HGDA suffers severe filter divergence with a very small ensemble size ( $k = 3$ ). Compared with global orthogonality, the reliance on local orthogonality degrades the hybrid system for small ensemble sizes (e.g.,  $k \leq 10$ ). However, the local QR-HGDA starts to show some improvement over the global one at  $k = 15$  and  $k = 20$ . This suggests that the local orthogonal component may alleviate the degradation obtained with the global orthogonalization. Furthermore, as discussed in the previous section, a well-tuned static  $\mathbf{B}$  matrix is needed to optimize the performance of the HGDA and the QR-HGDA. As shown in Table 4, the results of the global QR-HGDA and the local variant with the LETKF-derived  $\mathbf{B}$  matrix are consistent with those using the original  $\mathbf{B}$  matrix and the local QR-HGDA generally has worse performance than the global QR-HGDA.

TABLE 4. 50-day average RMSE for the QR-HGDA and the local QR-HGDA with different ensemble sizes ( $k$ ) using the LETKF-derived  $\mathbf{B}$  matrix. It can be compared with the cases with the original  $\mathbf{B}$  matrix shown in Table 1. Note the RMSEs are multiplied by a factor 100.

50-day avg RMSE ( $\times 100$ )	Global QR-HGDA	Local QR-HGDA	Difference
$k = 3$	2.6186	22.7854	20.1668
$k = 4$	1.2283	2.6725	1.4442
$k = 5$	0.9392	1.3406	0.4014
$k = 6$	0.8562	0.9902	0.1340
$k = 7$	0.7249	0.8721	0.1472
$k = 8$	0.7590	0.8867	0.1277
$k = 9$	0.6478	0.7197	0.0719
$k = 10$	0.6206	0.6571	0.0365
$k = 15$	0.5908	0.6086	0.0178
$k = 20$	0.5330	0.5289	-0.0041

The performance of the local QR-HGDA quickly improves with increasing ensemble size. The accuracy of two methods converges with large ensemble size and the local QR-HGDA is slightly more accurate than the global variant at  $k = 20$ .

### 7. Conclusions

The primary motivation of this study was to provide a method to eliminate the empirically determined hybrid weighting parameter used in conventional hybrid data assimilation systems. Based on the framework of the two-step hybrid gain data assimilation (HGDA) algorithm proposed by Penny (2014), the use of a hybrid weighting parameter can be avoided by limiting the variational correction to the subspace orthogonal to the linear subspace defined by the analysis ensemble perturbations. The orthogonal component is extracted by applying a QR factorization to the combined EnKF/VAR solution space. This orthogonal component is applied directly to update the EnKF analysis mean state rather than using a weighting parameter. This new algorithm is referred to as QR-HGDA.

The feasibility of the QR-HGDA algorithm is explored based on its performance with a quasigeostrophic (QG) model and compared to the parameter-dependent HGDA. By removing the dependency of the hybrid methods on a hybrid weighting parameter, we have highlighted the critical importance of carefully tuning the background ( $\mathbf{B}$ ) and observation ( $\mathbf{R}$ ) error covariance matrices. Given that the optimization of the hybrid gain matrix is the key characteristic of the hybrid gain methods, and that the Kalman gain matrix is a function of both  $\mathbf{B}$  and  $\mathbf{R}$ , we showed that poor estimation of either of these matrices can adversely affect the hybrid DA performance. This was shown for both the

QR-HGDA and HGDA based on their sensitivity to the  $\mathbf{B}$  and  $\mathbf{R}$  matrices and to model bias.

From a series of sensitivity tests, we found that a  $\mathbf{B}$  matrix well tuned for use with a stand-alone 3DVAR may not be the optimal choice for use in the hybrid methods. This highlights the imperative of evaluating the optimality of the climatological  $\mathbf{B}$  matrix not only in the HCDA (Satterfield et al. 2018), which combining the background error covariance matrix from component DA systems, but also in the HGDA. Degradation occurred in the hybrids because the 3DVAR-tuned  $\mathbf{B}$  matrix had a larger variance from large to middle spatial scales and quickly tapered off for small scales. With a sufficient ensemble size (e.g.,  $k \geq 9$ ), the LETKF already constrains the large scales fairly well. Thus for the HGDA, as long as the ensemble size is sufficient to stabilize the LETKF system, the high variance at larger spatial scales in the 3DVAR-tuned  $\mathbf{B}$  matrix used in the second step correction only causes the introduction of unnecessary corrections in the DA cycle that degrade accuracy. In contrast, the climatological  $\mathbf{B}$  derived from a long history of cycled LETKF forecasts has much smaller variance at large scales owing to the flow dependence in the LETKF corrections leading to forecast errors dominated by growing errors associated with local instabilities and partially to the use of localization. Compared to using the original climatological  $\mathbf{B}$  matrix, the accuracy of both the HGDA and QR-HGDA analyses are improved when instead using the LETKF-derived  $\mathbf{B}$  matrix.

When the observation error variance is intentionally overestimated by using perfect observations (i.e., without the presence of observational noise), the QR-HGDA has smaller RMSE than the HGDA and the standard LETKF. When observation noise is increased, the HGDA can partially cancel some of the noise induced by observations. It is noticed that the LETKF can ignore noise in directions that do not align with the growing error modes represented by the ensemble perturbations. Thus, the LETKF is less sensitive to the statistically isometric observational noise whereas the variational correction is more sensitive to those noises. If noise projects onto the orthogonal subspace relative to the analysis ensemble perturbation subspace, the QR-HGDA degrades the analysis accuracy more than the HGDA. This implies that the QR-HGDA is more sensitive to the performance of the VAR.

We would like to emphasize again that the LETKF-derived  $\mathbf{B}$  matrix is not the optimal one for using in the variational system. It shows a different error structure from the original  $\mathbf{B}$  matrix that is constructed using the forecast error collected from the stand-alone 3DVAR and thus contains both growing and nongrowing modes.

Given that the VAR component is used to correct the LETKF analysis mean, the need to reduce the large-scale error structure in the static  $\mathbf{B}$  matrix used in the second step HGDA raises the importance of reestimating the static  $\mathbf{B}$  matrix for the use with the HGDA. However, how to properly estimate a static background error is arduous, especially for an operational purpose. Even with this simple QG-model, it is a nontrivial task to optimize the static  $\mathbf{B}$  matrix. In real applications, a static  $\mathbf{B}$  matrix is usually preestimated and the only parameter occurred in a hybrid algorithm is the combination weight. When the background and observation error covariance matrices are well estimated, the QR-HGDA has better accuracy than the HGDA using the empirically best combination weighting. These results suggest that the QR-HGDA could be a competitive “parameterless” hybrid algorithm compared to other hybrid methods that rely on weighting parameters. Moreover, because of its greater sensitivity to the variational component, exploration with the QR-HGDA can help practitioners to evaluate and tune the background and observation error covariance matrices used in hybrid methods.

Finally, we explored whether the QR-HGDA can be performed with local orthogonality. For an operational center, it is not only the scalability of the DA system that is of interest but also the total computational cost. The HGDA and QR-HGDA provide excellent scalability of the hybrid DA system and the local QR-HGDA is advantageous to the computational cost. The findings with the QG model suggest that applying global orthogonality is more accurate with small ensemble sizes and the performance between the global QR-HGDA and the local QR-HGDA converges when the ensemble size increases. However, based on the experiments with the QG model, we found that global orthogonality is more advantageous for stabilizing the hybrid DA system and maintains large-scale balances. Furthermore, applying the variational correction with global orthogonality still has its importance in terms of constraining dynamical balance and large-scale bias.

In this work, the idea of limiting the variational correction to an orthogonal subspace to develop a parameterless hybrid algorithm is demonstrated with a simple dynamical model. The comparable accuracy of the local and the global QR-HGDA with sufficient ensemble size implies that local QR-HGDA has a potential for implementing a complex dynamical model with a higher dimension. While this finding is for a simple configuration, it can serve as a basis for further study of the QR-HGDA with full physics and dynamics to understand the feasibility in a realistic application. Moreover, developing an “online” algorithm that allows the

orthogonality procedure inside the variational system through the minimization process is worthy of further investigation. An additional interesting avenue of investigation might be to examine the impact of defining the orthogonality with a non-Euclidean inner product.

*Acknowledgments.* The authors are very grateful for the valuable comments from three anonymous reviewers, who have helped improve this manuscript. C.-C. Chang would like to acknowledge support from Taiwan Ministry of Science and Technology (MOST) Grants 105-2917-I-008 -004, 108-2119-M-002-022, and 108-2621-M-008 -003. Shu-Chih Yang is sponsored by the Taiwan Ministry of Science and Technology (MOST) Grant 108-2811-M-008 -583. S.G. Penny acknowledges support from the National Oceanic and Atmospheric Administration (NOAA) National Environmental Satellite, Data, and Information Service (NESDIS) [NA14NES4320003], the NOAA Climate Program Office (CPO) [NA16OAR4310140], the NOAA Next Generation Global Prediction System (NGGPS) program [NA18NWS4680048], the National Oceanographic Partnership Program (NOPP) supported by the Office of Naval Research (ONR), and the Indian Institute of Tropical Meteorology (IITM) Monsoon Mission II (MM-II) [IITMMIIUNIVMARYLANDUSA2018INT2].

#### REFERENCES

- Bannister, R. N., 2008: A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics. *Quart. J. Roy. Meteor. Soc.*, **134**, 1971–1996, <https://doi.org/10.1002/qj.340>.
- Barker, D. M., 1998: Var scientific development paper 25: The use of synoptic-dependent error structure in 3DVAR. U.K. Met Office Tech. Rep., 2 pp.
- Bishop, C. H., J. S. Whitaker, and L. Lei, 2017: Gain form of the ensemble transform Kalman filter and its relevance to satellite data assimilation with model space ensemble covariance localization. *Mon. Wea. Rev.*, **145**, 4575–4592, <https://doi.org/10.1175/MWR-D-17-0102.1>.
- Bocquet, M., and A. Carrassi, 2017: Four-dimensional ensemble variational data assimilation and the unstable subspace. *Tellus*, **69A**, 1304504, <https://doi.org/10.1080/16000870.2017.1304504>.
- Bonavita, M., M. Hamrud, and L. Isaksen, 2015: EnKF and hybrid gain ensemble data assimilation. Part II: EnKF and hybrid gain results. *Mon. Wea. Rev.*, **143**, 4865–4882, <https://doi.org/10.1175/MWR-D-15-0071.1>.
- Bormann, N., M. Bonavita, R. Dragani, R. Eresmaa, M. Matricardi, and A. McNally, 2016: Enhancing the impact of IASI observations through an updated observation-error covariance matrix. *Quart. J. Roy. Meteor. Soc.*, **142**, 1767–1780, <https://doi.org/10.1002/qj.2774>.
- Buehner, M., 2005: Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational NWP setting. *Quart. J. Roy. Meteor. Soc.*, **131**, 1013–1043, <https://doi.org/10.1256/qj.04.15>.
- , P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: One-month experiments with real observations. *Mon. Wea. Rev.*, **138**, 1567–1586, <https://doi.org/10.1175/2009MWR3158.1>.
- Carrassi, A., A. Trevisan, L. Descamps, O. Talagrand, and F. Uboldi, 2008: Controlling instabilities along a 3DVar analysis cycle by assimilating in the unstable subspace: A comparison with the EnKF. *Nonlinear Processes Geophys.*, **15**, 503–521, <https://doi.org/10.5194/npg-15-503-2008>.
- Chang, C.-C., S.-C. Yang, and C. Keppenne, 2014: Applications of the mean recentering scheme to improve typhoon track prediction: A case study of Typhoon Nanmadol (2011). *J. Meteor. Soc. Japan*, **92**, 559–584, <https://doi.org/10.2151/jmsj.2014-604>.
- Corazza, M., and Coauthors, 2003: Use of the breeding technique to estimate the structure of the analysis “errors of the day.” *Nonlinear Processes Geophys.*, **10**, 233–243, <https://doi.org/10.5194/npg-10-233-2003>.
- , E. Kalnay, and S. C. Yang, 2007: An implementation of the Local Ensemble Kalman Filter in a quasi geostrophic model and comparison with 3D-Var. *Nonlinear Processes Geophys.*, **14**, 89–101, <https://doi.org/10.5194/npg-14-89-2007>.
- Dee, D., 2005: Bias and data assimilation. *Quart. J. Roy. Meteor. Soc.*, **131**, 3323–3343, <https://doi.org/10.1256/qj.05.137>.
- Derber, J., and F. Bouttier, 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, **51A**, 195–221, <https://doi.org/10.3402/tellusa.v51i2.12316>.
- Etherton, B. J., and C. H. Bishop, 2004: Resilience of hybrid ensemble/3DVAR analysis schemes to model error and ensemble covariance error. *Mon. Wea. Rev.*, **132**, 1065–1080, [https://doi.org/10.1175/1520-0493\(2004\)132<1065:ROHDAS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<1065:ROHDAS>2.0.CO;2).
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10143, <https://doi.org/10.1029/94JC00572>.
- Golub, G. H., and C. F. Van Loan, 2013: *Matrix Computations*. 4th ed. Johns Hopkins University Press, 784 pp.
- Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter–3D variational analysis scheme. *Mon. Wea. Rev.*, **128**, 2905–2919, [https://doi.org/10.1175/1520-0493\(2000\)128<2905:AHEKFV>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<2905:AHEKFV>2.0.CO;2).
- Heemink, A. W., M. Verlaan, and A. J. Segers, 2001: Variance reduced ensemble Kalman filtering. *Mon. Wea. Rev.*, **129**, 1718–1728, [https://doi.org/10.1175/1520-0493\(2001\)129<1718:VREKF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<1718:VREKF>2.0.CO;2).
- Houtekamer, P. L., and F. Zhang, 2016: Review of the ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **144**, 4489–4532, <https://doi.org/10.1175/MWR-D-15-0440.1>.
- , M. Buehner, and M. De La Chevrotière, 2018: Using the hybrid gain algorithm to sample data assimilation uncertainty. *Quart. J. Roy. Meteor. Soc.*, **145**, 35–56, <https://doi.org/10.1002/qj.3426>.
- Huang, B., and X. Wang, 2018: On the use of cost-effective valid-time-shifting (VTS) method to increase ensemble size in the GFS hybrid 4DVar system. *Mon. Wea. Rev.*, **146**, 2973–2998, <https://doi.org/10.1175/MWR-D-18-0009.1>.
- Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126, <https://doi.org/10.1016/j.physd.2006.11.008>.
- Isaksen, L., M. Bonavita, R. Buizza, M. Fisher, J. Haseler, M. Leutbecher, and L. Raynaud, 2010: Ensemble of data

- assimilations at ECMWF. ECMWF Tech. Memo. 636, 45 pp., [http://old.ecmwf.int/publications/library/ecpublications/\\_pdf/tm/601-700/tm636.pdf](http://old.ecmwf.int/publications/library/ecpublications/_pdf/tm/601-700/tm636.pdf).
- Kalnay, E., and S.-C. Yang, 2010: Accelerating the spin-up of ensemble Kalman filtering. *Quart. J. Roy. Meteor. Soc.*, **136**, 1644–1651, <https://doi.org/10.1002/qj.652>.
- Kleist, D., 2012: An evaluation of hybrid variational-ensemble data assimilation for the NCEP GFS. Ph.D. dissertation, University of Maryland, 149 pp.
- , and K. Ide, 2015: An OSSE-based evaluation of hybrid variational-ensemble data assimilation for the NCEP GFS. Part I: System description and 3D-hybrid results. *Mon. Wea. Rev.*, **143**, 433–451, <https://doi.org/10.1175/MWR-D-13-00351.1>.
- Leon, S. J., Å. Björck, and W. Gander, 2013: Gram–Schmidt orthogonalization: 100 years and more. *Numer. Linear Algebra Appl.*, **20**, 492–532, <https://doi.org/10.1002/nla.1839>.
- Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP—A comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183–3203, <https://doi.org/10.1256/qj.02.132>.
- Ménétrier, B., and T. Auligné, 2015: Optimized localization and hybridization to filter ensemble-based covariances. *Mon. Wea. Rev.*, **143**, 3931–3947, <https://doi.org/10.1175/MWR-D-15-0057.1>.
- Mitchell, H. L., and P. L. Houtekamer, 2000: An adaptive ensemble Kalman filter. *Mon. Wea. Rev.*, **128**, 416–433, [https://doi.org/10.1175/1520-0493\(2000\)128<0416:AAEKF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<0416:AAEKF>2.0.CO;2).
- Morss, R. E., 1999: Adaptive observations: Idealized sampling strategies for improving numerical weather prediction. Ph.D. thesis, Massachusetts Institute of Technology, 255 pp.
- Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763, [https://doi.org/10.1175/1520-0493\(1992\)120<1747:TNMCSS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2).
- Penny, S. G., 2014: The hybrid local ensemble transform Kalman filter. *Mon. Wea. Rev.*, **142**, 2139–2149, <https://doi.org/10.1175/MWR-D-13-00131.1>.
- , 2017: Mathematical foundations of hybrid data assimilation from a synchronization perspective. *Chaos*, **27**, 126801, <https://doi.org/10.1063/1.5001819>.
- , D. Behringer, J. Carton, and E. Kalnay, 2015: A hybrid global ocean data assimilation system at NCEP. *Mon. Wea. Rev.*, **143**, 4660–4677, <https://doi.org/10.1175/MWR-D-14-00376.1>.
- Petrie, R. E., and R. N. Bannister, 2011: A method for merging flow-dependent forecast error statistics from an ensemble with static statistics for use in high-resolution variational data assimilation. *Comput. Fluids*, **46**, 387–391, <https://doi.org/10.1016/j.compfluid.2011.01.037>.
- Rotunno, R., and J.-W. Bao, 1996: A case study of cyclogenesis using a model hierarchy. *Mon. Wea. Rev.*, **124**, 1051–1066, [https://doi.org/10.1175/1520-0493\(1996\)124<1051:ACSOCU>2.0.CO;2](https://doi.org/10.1175/1520-0493(1996)124<1051:ACSOCU>2.0.CO;2).
- Satterfield, E. A., D. Hodyss, D. D. Kuhl, and C. H. Bishop, 2018: Observation-informed generalized hybrid error covariance models. *Mon. Wea. Rev.*, **146**, 3605–3622, <https://doi.org/10.1175/MWR-D-18-0016.1>.
- Snyder, C., T. M. Hamill, and S. B. Trier, 2003: Linear evolution of error covariances in a quasigeostrophic model. *Mon. Wea. Rev.*, **131**, 189–205, [https://doi.org/10.1175/1520-0493\(2003\)131<0189:LEOECI>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<0189:LEOECI>2.0.CO;2).
- Storto, A., P. Oddo, A. Cipollone, I. Mirouze, and B. Lemieux, 2018: Extending an oceanographic variational scheme to allow for affordable hybrid and four-dimensional data assimilation. *Ocean Modell.*, **128**, 67–86, <https://doi.org/10.1016/j.oceanmod.2018.06.005>.
- Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317–2330, [https://doi.org/10.1175/1520-0477\(1993\)074<2317:EFANTG>2.0.CO;2](https://doi.org/10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2).
- Verlaan, M., and A. W. Heemink, 1997: Tidal flow forecasting using reduced rank square root filters. *Stoch. Hydrol. Hydraul.*, **11**, 349–368, <https://doi.org/10.1007/BF02427924>.
- Wang, X., D. F. Parrish, D. T. Kleist, and J. S. Whitaker, 2013: GSI 3DVAR-based ensemble-variational hybrid data assimilation for NCEP Global Forecast System: Single-resolution experiments. *Mon. Wea. Rev.*, **141**, 4098–4117, <https://doi.org/10.1175/MWR-D-12-00141.1>.
- Yang, S.-C., M. Corazza, A. Carrassi, E. Kalnay, and T. Miyoshi, 2009a: Comparison of local ensemble transform Kalman filter, 3DVAR, and 4DVAR in a quasigeostrophic model. *Mon. Wea. Rev.*, **137**, 693–709, <https://doi.org/10.1175/2008MWR2396.1>.
- , E. Kalnay, B. Hunt, and N. E. Bowler, 2009b: Weight interpolation for efficient data assimilation with the local ensemble transform Kalman filter. *Quart. J. Roy. Meteor. Soc.*, **135**, 251–262, <https://doi.org/10.1002/qj.353>.
- , —, and —, 2012a: Handling nonlinearity in an ensemble Kalman filter: Experiments with the three-variable Lorenz model. *Mon. Wea. Rev.*, **140**, 2628–2646, <https://doi.org/10.1175/MWR-D-11-00313.1>.
- , —, and T. Miyoshi, 2012b: Accelerating the EnKF spinup for typhoon assimilation and prediction. *Wea. Forecasting*, **27**, 878–897, <https://doi.org/10.1175/WAF-D-11-00153.1>.
- , —, and T. Enomoto, 2015: Ensemble singular vectors and their use as additive inflation in EnKF. *Tellus*, **67A**, 26536, <https://doi.org/10.3402/tellusa.v67.26536>.